A THESIS FOR THE DEGREE OF MASTER OF SCIENCE

Molecular Breeding of Pepper Varieties (*Capsicum annuum*) Containing High Levels of Capsinoid

캡시에이트 고함유 고추 (*Capsicum annuum*) 계통의 분자육종

AUGUST, 2015

HYEONSEOK JEONG

MAJOR IN HORTICULTURAL SCIENCE
DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

Molecular Breeding of Pepper Varieties (*Capsicum annuum*) Containing High Levels of Capsinoid

UNDER THE DIRECTION OF DR. BYOUNG-CHEORL KANG
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

BY
HYEONSEOK JEONG

MAJOR IN HORTICULTURAL SCIENCE
DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

AUGUST, 2015

APPROVED AS A QUALIFIED DISSERTATION OF HYEONSEOK JEONG
FOR THE DEGREE OF MASTER OF SCEINCE
BY THE COMMITTEE MEMBERS

CHAIRMAN _____
Jin Hoe Huh, Ph.D.

VICE-CHAIRMAN _____
Byoung-Cheorl Kang, Ph.D.

MEMBER _____
Doil Choi, Ph.D.

Molecular Breeding of Pepper Varieties (*Capsicum annuum*) Containing High Levels of Capsinoid

HYEONSEOK JEONG

DEPARTMENT OF PLANT SCIENCE

THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

ABSTRACTS

Capsinoid, low-pungent compounds, have the same biological effects as capsaicinoid such as anticancer and anti-obesity. A precursor of capsinoid, vanillyl alcohol, is known to be produced by mutations in the *putative-aminotransferase* (*pAMT*) gene. In the previous study, 'SNU11-001' (*Capsicum chinense*) containing high levels of capsinoid was identified in germplasm collections of *Capsicum*. This collection has a unique mutation in the *pAMT* gene that can cause dysfunction of this gene. In order to develop pepper varieties containing a high capsinoid content, marker-assisted foreground and background selections were performed during backcross breeding. Compared to the conventional backcrossing, marker-assisted

backcrossing (MABC) is extremely useful for recovery of a recurrent parent's genetic background. For foreground selection, plants carrying the *pAMT/pamt* genotype were selected from a BC_1F_1 and BC_2F_1 populations using SCAR markers derived from the unique *pAMT* mutation of 'SNU11-001'. To obtain background selection markers, a total of 412 single nucleotide polymorphism (SNP) markers was screened on 'Shinhong' parental lines and 'SNU11-001' to obtain polymorphic SNP markers. Of the 412 SNP markers, 144 and 204 polymorphic SNP markers evenly distributed in pepper genome were finally selected. BC_1F_1 and BC_2F_1 plants carrying the *pAMT/pamt* genotype were subjected to background selection using the selected marker sets. Multiple genotype analysis was done using a high-throughput genotyping system (EP1™, Fluidigm®, USA). As a result, one BC_1F_1 plant 84% similar to the recurrent parent and several BC_2F_1 plants more than 96% recovery rate of the recurrent parent were selected. Genetic backgrounds of the selected BC_2F_1 plants were evaluated by the genotype-by-sequencing (GBS) method in order to confirm the background selection results using the SNP marker set. GBS results showed that recovery rate and positions of introgressed segments were well matched between two methods demonstrating MABC can be successfully done with a couple hundred SNP markers.

Keywords: Capsinoid, *pAMT*, marker-assisted backcross breeding

Student number: 2013-23232

CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CS	Capsaicin synthase
CMS	Cytoplasmic male sterility
CTAB	Cetyl trimethylammonium bromide
GBS	Genotyping by sequencing
HPLC	High-performance liquid chromatography
MABC	Marker-Assisted backcross breeding
pAMT	*putative-Aminotransferase*
Rf	*Fertility restorer*
SCAR	Sequence characterized amplified regions
SNP	Single nucleotide polymorphism
Tcc	Transposon of *C. chinense*

INTRODUCTION

The unique characteristic of pepper is pungency, which is caused by capsaicinoid in fruits. Capsaicin, a member of the capsaicinoid, is produced by capsaicin synthase (CS) through condensation of branched-chain fatty acids and vanillylamine produced from vanillin by *p-AMT* (Curry et al., 1999). Capsaicinoids have many biomedical functions such as anticancer and anti-obesity (Luo et al., 2011; Thiele et al., 2008). But capsaicinoid sometimes causes inflammation and fever due to its pungency (Tanaka et al., 2010a).

Pepper cultivar 'CH-19 sweet' containing capsaicinoid-like substances were first reported by Yazawa et al. (1989). After a decade of research, it was revealed that molecular structures of the compounds are similar to capsaicinoid, and these substances were named as capsiate and dihydrocapsiate (Kobata et al., 1998). Capsinoid, capsaicinoid analogues, has a less-pungent character unlike capsaicinoid (Sasahara et al., 2010). Molecular structures of capsinoid and capsaicinoid are similar but capsinoid has an ester bond at the place of an amide bond (Tanaka et al., 2010a). Capsinoid shows similar biological functions to capsaicinoid including antitumor, antioxidant and anti-obesity. Due to its low pungency, capsinoid has advantages over capsaicinoid in clinical applications (Luo et al., 2011).

Biosynthesis of capsinoid has not been clearly revealed. In the hot pepper vanillin is converted to vanillyamine and vanillyl alcohol both (Kobata et al., 2013).

It is known that dysfunctional mutations of *pAMT* result in impediment of the formation vanillylamine from vanillin (Lang et al., 2009; Tanaka et al., 2010a). In the *pamt* mutant, vanillin was converted exclusively to vanillyl alcohol instead of vanillylamine, and then vanillyl alcohol is utilized as one of substrates for capsiate biosynthesis by capsaicin synthase (CS) (Han et al., 2012; Tanaka et al., 2010b). Reported *pamt* mutations until now are nonsense mutation in 'CH19-Sweet' (Lang et al., 2009), transversion mutation, cysteine to arginine, in 'Himo' (Tanaka et al., 2010a), frameshift mutations in 'Belize Sweet', 'Zavory hot' and 'Aji dulce (strain 2)' (Tanaka et al., 2010b) and 12-bp deletion mutation in 'S3212' (Park et al., 2015).'CH-19 Sweet' and non-pungent pepper 'Murasaki' were crossed for developing low-pungent peppers containing capsinoid. As a result, 'Maru Salad' containing 700 μg/gDW was developed (Tanaka et al., 2014). Jang et al. (2015) showed that substitution of a dysfunctional *pAMT* allele results in low-pungency but high levels of capsicinoids in *C. chinense* 'Habanero'. This results demonstrate that the genetic factors controlling the level of capsaicinoid also control that of capsinoid..

Backcross breeding was first introduced in 1922 and has been used for an effective breeding method to introgress one or a few genes to an elite cultivar (Stoskopf et al., 1993). Backcross breeding was mostly used for introgression of single Mendelian traits like a disease resistance (Johnston, 1974). However, conventional backcross breeding is a time consuming process. Marker-assisted backcrossing (MABC) is a method to expedite the backcross breeding process by

selecting individuals containing high recurrent parent background using molecular markers (Hospital and Charcosset, 1997). MABC is consisted of two steps, 'foreground selection' and 'background selection'. 'Foreground selection' is a step for selecting individuals harboring genes related with target traits in the backcross population by using molecular markers. It is more effective when identification of individual plants with target traits is difficult or labor consuming (Tanksley, 1983). 'Background selection' is a step for selecting individuals with closest to the recurrent parent's gentic background using molecular markers (Hospital et al., 1992). MABC can greatly reduce breeding time as compared to the conventional backcross breeding. For instance, MABC approach greatly shortened breeding time in wheat (Randhawa et al., 2009).

In this study, we developed MABC program for a development of new pepper cultivars containing capsinoid. Towards this end, 'SNU11-001' was used as a donor for the dysfunctional *pAMT* and 'Shinhong' C lines was used as a recurrent parent. Individual plants having the heterozygous genotype for *pAMT* marker were selected in BC_1 and BC_2 populations and then plants showing the most recovered 'Shinhong C' genetic background were selected by a set of SNP markers evenly distributed in pepper genome. As a result, candidates for 'Shinhong C' harboring the dysfunctional *pAMT* could be developed in two backcross generations.

LITERATURE REVIEW

1. Biochemistry of capsinoid

Capsaicinoid-like compounds were first reported in the pepper cultivar 'CH-19 Sweet' by Yazawa in 1989 and after a decade, that compounds were named capsiate, dihydrocapsiate and nordihydrocapsiate and that were grouped to capsinoid (Kobata et al., 1999; Kobata et al., 1998). Structure of capsinoids are similar to capsaicinoids, but a peptide bond is replaced by an ester bond (Kobata et al., 1999; Kobata et al., 1998). Capsiate is less stable than capsaicin and decomposed easily in aqueous phase while capsaicinoid are stable in nonpolar as well as polar solvent condition (Sutoh et al., 2001).

2. Biomedical effect of capsinoid

Clinical benefit of capsinoid has not been reported as much as capsaicinoid but it is more attractive for pharmaceuticals because of less-pungent character than capsaicinoid (Sasahara et al., 2010). Iida reported that similar to capsaicinoids, capsinoids were also able to activate transient receptor potential vanilloid subfamily member 1 (TRPV1) and prevent pain perception (Iida et al., 2003). Capsinoid also have anticancer activity like capsaicinoid by exhibited potent activity against vascular endothelial growth factor (VEGF)-induced angiogenesis and vascular permeability (Pyun et al., 2008). Capsinoid is effective to reduce weight by

enhancing energy expenditure (Faraut et al., 2009), activating sympathetic nervous system and raising body temperature (Kawabata et al., 2009).

3. Biosynthesis of capsinoid

Capsaicin is synthesized by condensation of vanillylamine and 8-methyl-6-nonenoic acid. *pAMT* converts vanillin into vanillylamine in phenylpropanoid pathway and *CS* produces capsaicin using vanillylamine and branched fatty acid, 8-methyl-6-nonenoic acid, in the final step of capsaicin biosynthesis. On the other hand, *pAMT* mutation causes formation of vanillyl alcohol instead of vanillylamine from vanillin. Capsiate is synthesized from vanillyl alcohol and branched chain fatty acid, 8-methyl-5-nonenoyl-CoA (Tanaka et al., 2010a).

It is quite recent that *pAMT* mutation was reported to have influence on capsinoid biosynthesis. *C. annuum* 'CH-19 Sweet' containing high levels of capsinoid has a frameshift mutation owing to T insertion on *pAMT* but there is no significant difference in its transcription level compared to pungent cultivar (Lang et al., 2009). *C. annuum* 'Himo' also produce capsinoid. An amino acid substitution from cystein to arginine resulted in dysfunctional *pAMT*. Single amino acid substitution was involved in pyridoxal 5-phosphate (PLP) binding domain that influences on *pAMT* enzyme activity (Tanaka et al., 2010a). A transposable element was found in this gene in *C. chinense* cultivars (Zavory hot, Aji Dulce stain 2) with high content of capsinoid (Tanaka et al., 2010b). ‘Transposable element of *C. chinense* (*Tcc*)’ is consisted of approximately 2.3 kb nucleotide sequence. *Tcc* was

inserted in the fifth intron of Zavory hot and the third intron of Aji Dulce stain 2. *Tcc* affected synthesis of functional mRNA of *pAMT*. *Pun1* was also elucidated to control biosynthesis of capsinoid (Han et al., 2012). An F_6 recombinant inbred line population derived from SR211, pungent pepper with capsiniods and SR213, nonpungent pepper was used to determine the role of *Pun1* in capsinoid production. Plants with *Pun1*/*Pun1* genotype could synthesize both capsaicinoid and capsinoid but plants with *pun1*/*pun1* could not produce capsaicinoid as well as capsinoid. Therefore functional *Pun1* is necessary to produce capsinoid and nonfunctional *pAMT* is necessary condition for biosynthesis of capsinoid.

4. Marker-assisted backcross breeding

Marker-assisted backcrossing (MABC) is a backcross program selecting individuals containing high recurrent parent background using molecular markers (Hospital and Charcosset, 1997). The first application of molecular markers to the backcross program was unclear. In 1980, Tanksley predicted the recovery rate to the recurrent parent according to the number of marker (Tanksley and Rick, 1980). In 1995, Ragot applied RFLP markers to the maize backcross program to introgress phosphinothricin resistance and insecticidal protein genes (Ragot et al., 1995). As a result of that study, 99.36% recovered to the recurrent parent genotype plant was obtained. The effectiveness of MABC depends on the selection strategies was studied. Frisch carried out simulations with a maize genetic map consisting of 80 markers to investigate effectiveness of MABC according to the population size

(Frisch et al., 1999). He claimed that increasing population sizes from BC_1 to BC_3 generation can reduce the number of required marker data points by as much as 50%. Herzog and Frisch also studied effective MABC strategies by using computer simulations and suggested several strategies that marker density of 10 cM for background selection is sufficient and adding the recombinant selection step using flanking markers to target gene is more effective rather than only the foreground selection (Herzog and Frisch, 2011).

MATERIALS AND METHODS

Plant materials and backcross population

C. annuum 'Shinhong A, B, C' lines and *C. chinense* 'SNU11-001' were used for the parental lines in this study. 'Shinhong A' has a sterile cytoplasm (*S*), the recessive homozygous *restorer-of-fertility* (*Rf*) gene and the dominant homozygous *pAMT* gene. 'Shinhong B' (*C. annuum*) has a normal cytoplasm (*N*), the recessive homozygous *Rf* gene and the dominant homozygous *pAMT* gene. 'Shinhong B' is the maintainer of 'Shinhong A'. 'Shinhong C' (*C. annuum*) is a Korean landrace Jeju (Kang et al., 2014). 'Jeju' and has dominant homozygous *pAMT* gene. 'SNU11-001' is from Venezuela (Kang et al., 2014) and has the recessive homozygous *pAMT* gene.

Measurement of the capsaicinoid and capsinoid levels in fruits

Mature fruits excluding seeds were used for measuring the capsaicinoid and capsinoid content. HPLC analysis was performed in National Instrumentation Center for Environmental Management (NICEM, Seoul, Korea) according to the method described by Han et al. (2012).

Genomic DNA extraction

Total DNA was obtained from fully expanded cotyledon tissues of the each

plant using modified cetyl trimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987)

After DNA extraction, each sample was moved to 96 well plates. DNA concentrations were measured by NanoDrop (NanoDrop Technologies, Inc., Wilmington, DE, USA). DNA was diluted with 0.1 M TE buffer (PH 7.0) to final concentration of 10 ng/μl for SNP marker analysis. DNA samples were stored at -20℃ prior to use.

Development of the *pAMT* SCAR marker

C. annuum pAMT mRNA sequence was obtained from NCBI nucleotide database (GenBank: HM991733.1). Genomic DNA sequence was obtained from the blast of *pAMT* mRNA sequence against CM334 20140106 (v1.5) SCAFFOLD database (http://peppergenome.snu.ac.kr). *Tcc* sequence in 3rd intron of *pAMT* gene of *C. chinense* was obtained from published paper of Tanaka et al. (2010b). Two forward primers and one reverse primer were designed from Primer3web (v4.0.0, http://primer3.ut.ee/) using 3rd intron and 3rd exon of *pAMT* gDNA sequence and *Tcc* sequence.

Polymorphism survey of SNP markers between parental lines

‘Shinhong B’, ‘Shinhong C’ and ‘SNU11-001’ genomic DNA were used for

polymorphism survey. Genomic DNA was diluted with TDW to final concentration of 60 ng/μl. A total of 412 SNP markers developed from the previous study (Kang et al. (2014) for the Fluidigm® EP1^{TM} system were used. Genotyping was performed with the EP1^{TM} system and 96.96 Dynamic Array Chip for Genotyping (BMK-M-96.96GT) as described in the manufacturer's protocol.

pAMT genotyping

pAMT genotyping was performed with the combination of the primer set consisted of 'Capsiate_3rd_intron(F)', 'Capsiate_3rd_intron_R(LA)' and 'Tcc-R3'. PCR screening was conducted in a 25 μl reaction volume containing 100 ng of template DNA, 10 pmol of each primer, 2 μl of dNTP, 2.5 μl of Hipi buffer and 1 unit of Taq polymerase. PCR condition was as follows: 95℃ for 5 minutes followed by 35 cycles of 95℃ for 30 seconds, 58℃ for 30 seconds and 72℃ for 50 seconds and final extension of 10 minutes at 72℃.

Genome-wide genotyping

Genome-wide genotyping was performed with polymorphic markers between 'Shinhong C' and 'SNU11-001'. 192 SNP markers were used in BC_1 generation consiting of of 141 plants and 64 SNP markers were used in BC_2 generation of 89 plants. All plants used in this experiment had the heterozygous genotype for *pAMT*.

Genomic DNA concentration and genotyping method are the same as the process of the polymorphism survey. Recovery rate of the 'Shinhong C' was calculated as the percentage of Shinhong C alleles (the number of plants with 'Shinhong C' genotype + half of number of plants withheteozygous genotype)/total number of plants). A Graphical illustration of the selected plants was drawn using MapChart (v2.2, Wageningen UR, Netherlands) developed by Voorrips (2002).

GBS library preparation and sequencing

'SNU11-001', 'Shinhong C', one BC_1F_1 and seven BC_2F_1 plants which are selected on the basis of the top recovery rate were used in the GBS. 'SNU11-001' and 'Shinhong C' were tested in twice. All genomic DNA concentrations were 80 ng/μl. *PstI -MseI* sequencing libraries were constructed using SBG 100-Kit (v2.0, KeyGene N.V., Wageningen, Netherlands) as described in the manufacturer's protocol. After the construction of library the sequencing was performed using the Illumina HiSeq2000 at Macrogen (Macrogen, Seoul, Korea).

Verification of the recovery rate with GBS

Quality control (QC) and trimming for GBS Sequence data was performed using the CLC Genomics Workbench (v6.5, CLC bio, QIAGEN). Quality trimming set value was 0.01 (=Q20). Several software tools as Burrow Wheelers Alignment (BWA), Sequence Alignment/Map (SAM) Tools, and Picard were used for the

alignment of the GBS data. Sorting, read grouping, SNP calling and filtering were carried out using GenomeAnalysisTK (GATK). SNPs polymorphic between ‘Shinhong C’ and ‘SNU11-001’ were selected.

RESULTS

Characteristics parental lines for MABC

Plants and fruits shape used in this study are shown in Fig. 1 and Table 1. The plant habit of ‘Shinhong B’ and ‘Shinhong C’ are straight and ‘SNU11-001’ is spreading (Fig. 1A). Fruits of ‘Shinhong B’ are small and oval shape (Fig. 1B) whereas those of ‘Shinhong C’ are long and bent. ‘SNU11-001’ has short, round and crumpled shape fruits. Fruits of ‘Shinhong C’ are longer in length than those of ‘Shinhong B’ and ‘SNU11-001’. The fruit width of ‘SNU11-001’ is the largest whereas those of ‘Shinhong B’ is the smallest.

Capsaicinoid content in ‘Shinhong B’ was 7064.18±731.67 μg/gDW and that in ‘Shinhong C’ was 270.02±50.18 μg/gDW (Table 2). Capsaicinoid was not detected in ‘SNU11-001’. Capsinoid content in ‘Shinhong B’ was 164.48±82.51 μg/gDW and in ‘Shinhong C’ was 40.63±32.55 μg/gDW. Capsinoid content of SNU 001 was 9649.92±0 μg/gDW.

Table 1. The fruit characteristics of plants used in this study

Species	Cultivar	Fruit length (cm)	Fruit width (cm)	Fruit weight (g)
C. annuum	Shinhong B	2.75±0.38	1.66±0.14	1.35±0.47
C. annuum	Shinhong C	10.36±2.07	2.04±0.31	5.94±2.56
C. chinense	SNU11-001	2.78±0.29	3.68±0.15	4.18±0.34

Table 2. Capsaicinoid and capsinoid contents and *pAMT* genotypes of parental lines

Cultivar	*pAMT* genotype	Capsaicinoid (μg/gDW)[a]			Capsinoid (μg/gDW)		
		Capsaicin	Dihydrocapsaicin	Total	Capsiate	Dihydrocapsiate	Total
Shinhong B	*pAMT/pAMT*	4258.47±1185.47	2805.71±731.67	7064.18±731.67	147.71±38.45	22.37±5.27	164.48±82.51
Shinhong C	*pAMT/pAMT*	179.66±143.13	90.35±50.18	270.02±50.18	36.36±28.17	6.41±3.37	40.63±32.55
SNU11-001	*pamt/pamt*	nd[b]	nd	nd	8543.01±0	1106.91±0	9649.92±0

[a]DW = dry weight of the whole fruit excluding seeds.
[b]nd = not detected.

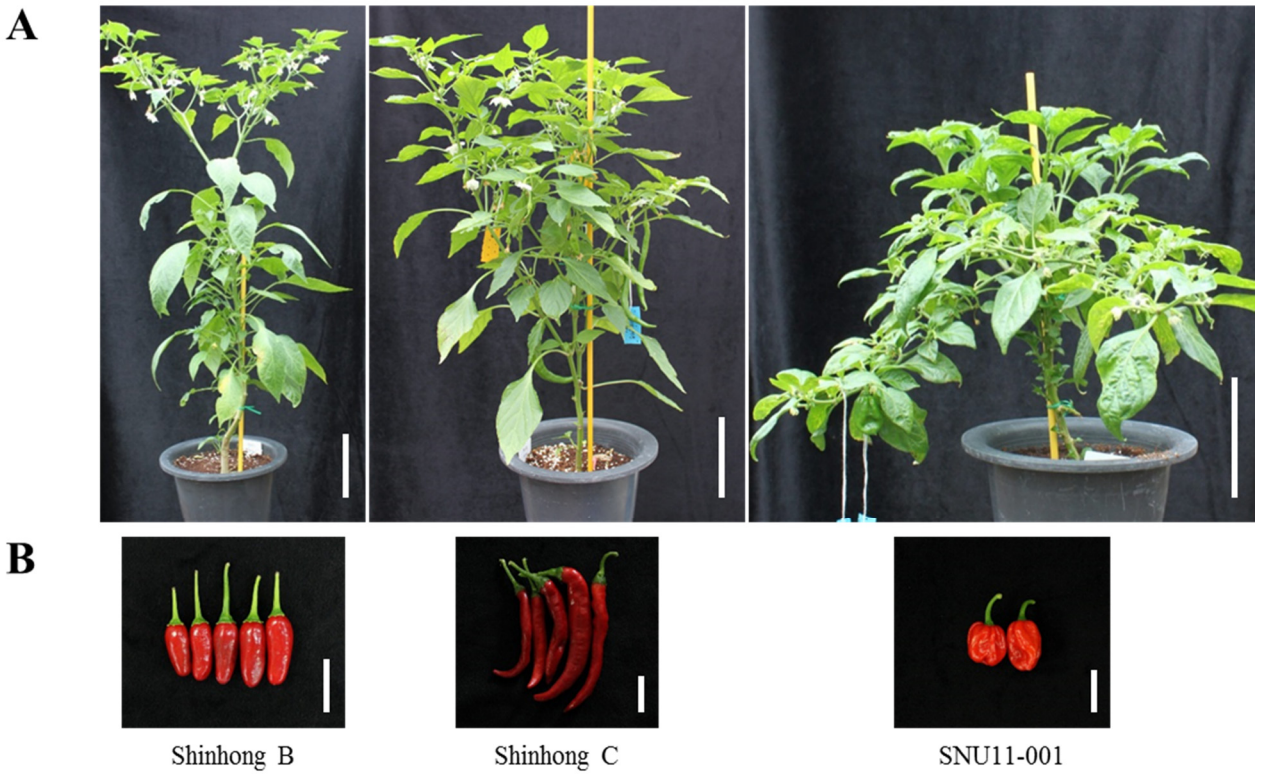

Fig. 1. ‘Shinhong’ parental lines and their fruits. ‘Shinhong B’, ‘Shinhong C’ and ‘SNU11-001’ is ordered from left to right. **A** Plants of ‘Sweet Shinhong’ parental lines. *Bar* 11 cm. **B** Fruits of the ‘Sweet Shinhong’ parental lines. *Bar* 3 cm.

Development of DNA markers for the MABC

1.1 Development of the *pAMT* SCAR marker for the foreground selection

To develop a SCAR marker for the *pAMT* gene, *Tcc* sequence was used. To find consensus and polymorphic sequences between 'SNU11-001' and 'Shinhong' parental lines, the 3rd intron and the 4th exon regions of the *pAMT* gene were sequenced. The 3rd-intron(F) primer was designed using consensus sequence of 3rd intron of *pAMT* and the 3rd-intron R(LA) primer was designed based on consensus sequence of the 4th exon (Fig. 2A). Tcc-R3 primer was designed using *Tcc* that is specific sequence in 'SNU11-001'. The expected PCR product size using the 3rd-intron(F) and 3rd-intron R(LA) primer pair is 619 bp in length predicted from scaffold339 (CM334 20140106 (v1.5) SCAFFOLD, http://peppergenome.snu.ac.kr) whereas those amplified by 3rd-intron(F) and 3rd intron(F) is 252 bp in length, which was estimated from Tanaka et al. (2010b).

PCR products of 'Shinhong B' and 'Shinhong C' were located at near 600 bp whereas that of 'SNU11-001' was located at 250 bp position (Fig. 2B) as expected. F_1 plants derived from the cross between 'Shinhong' parental lines and 'SNU11-001' showed both 619and 252 bp PCR products. BC_1 and BC_2 plants obtained from the cross had either 619 bp PCR product or both 619 and 252 bp PCR products.

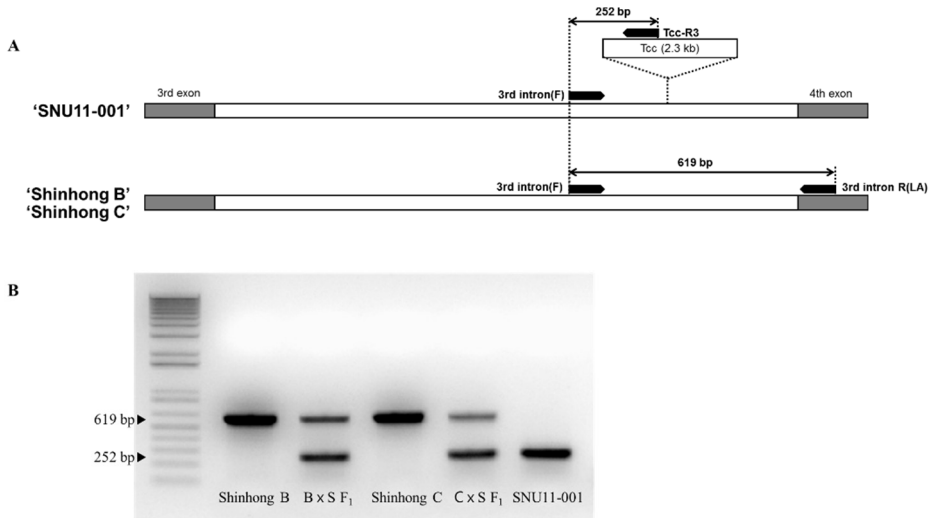

Fig. 2. Development of a SCAR marker to distinguish between *pAMT* and *pAMT* mutant. **A** *pAMT* SCAR marker primers location on the *pAMT* 3rd intron and 4th exon. One forward primer (3rd-intron(F)) and two reverse primer (Tcc-R3 and 3rd-intron R(LA)) were used for the marker analysis. 3rd-intron(F) is on the 3rd intron and Tcc-R3 is on the 2.3 kbp *Tcc* sequence of 'SNU11-001' and 3rd-intron R(LA) is on the 4th exon of the *pAMT* gene. The combination of 3rd intron(F) and 3rd intron R(LA) not worked in 'SNU11-001' gDNA because of the 2.3 kb *Tcc*. **B** Patterns of the *pAMT* SCAR marker PCR band. 'Shinhong' B and C showed 619 bp band size and 'SNU11-001' showed 252 bp band size. *B x S* 'Shinhong B' x 'SNU11-001' *C x S* 'Shinhong C' x 'SNU11-001'.

1.2 Polymorphism survey of the SNP markers between parental lines for the background selection

To select SNP markers polymorphic between 'Shinhong' parental lines and 'SNU11-001', a total of 412 SNP markers were analyzed using a high throughput SNP genotyping system (EP1^{TM}, Fluidigm, USA) (Table 3). These SNP markers are distributed in whole *Capsicum* chromosomes and an average marker density is about 3.55 cM being about 34 markers per chromosome. Chromosome with the most number of markers is chromosome 1 (69 markers) and the least number is chromosome 8 (8 markers).

When polymorphic markers were surveyed, a total of 144 and 204 SNP markers were polymorphic between 'Shinhong B' and 'SNU11-001' 'Shinhong, C' and 'SNU11-001', respectively (Table 4). Average polymorphic marker densities were 9.37 cM and 6.52 cM for 'Shinhong B' and 'SNU11-001' and 'Shinhong C' popluation, respectively. Number of polymorphic markers and marker densities according to chromosomes are summarized in Table 4. There were only two polymorphic markers on chromosome 8 for both populations.

Table 3. SNP markers for background selection

Chromosome	Number of SNP marker	Chromosome size (cM)[a]	Marker density (cM)[b]
1	69	177.1	2.57
2	43	113.8	2.65
3	57	141.6	2.48
4	32	144.1	4.50
5	25	101.6	4.06
6	36	136.6	3.79
7	21	104.2	4.96
8	8	50.3	6.29
9	22	110.2	5.01
10	35	98.0	2.80
11	33	90.0	2.73
12	31	106.3	3.43
Total	412	1460.6	3.55

[a] Chromosome size was calculated as the sum of the genetic distance between markers.
[b] Marker density was calculated by dividing chromosome size with the number of SNP marker.

Table 4. Number of polymorphic markers and marker density

Chromosome	'SNU11-001' x 'Shinhong B'		'SNU11-001' x 'Shinhong C'	
	Polymorphic marker	Marker density (cM)[a]	Polymorphic marker	Marker density (cM)
1	26	7.03	34	5.25
2	12	7.58	21	4.97
3	18	7.73	32	4.24
4	10	10.73	12	8.68
5	9	5.25	12	6.97
6	16	8.65	21	6.00
7	9	12.92	12	9.15
8	2	13.43	2	6.15
9	10	9.68	12	6.46
10	12	7.83	16	6.30
11	13	7.02	17	5.49
12	7	14.57	13	8.59
Total	144	9.37	204	6.52

[a] Marker density was calculated by dividing chromosome size with the number of polymorphic marker.

Marker-assisted backcrossing for ‘Shinhong’ C line containing the mutated *pAMT* allele

1.1 Foreground selection using *pAMT* marker

To select plants carrying the *pAMT* allele (*pAMT*/*pamt*) from the backcross populations, plants were subjected to genotype analysis using the *pAMT* marker (Fig. 3). A total of 250 BC_1F_1 plants derived from ‘Shinhong C’ x ‘SNU11-001’ were used for the first round foreground selection. As a result, 109 plants showed the homozygous genotype (*pAMT*/*pAMT*) for the *pAMT* allele whereas 141 plants showed the heterozygous genotype (*pAMT*/*pamt*) (Table 5). The plants carrying the heterozygous *pAMT* locus were used for background selection. A total of 160 BC_2F_1 from ‘Shinhong C’ x ‘SNU11-001’ were also used for the second round foreground selection. Among the BC_2F_1 plants, 71 plants showed the *pAMT*/*pAMT* genotype whereas 89 showed the heterozygous genotype (*pAMT*/*pamt*). Again, 89 having the heterozygous genotype were subjected for background selection.

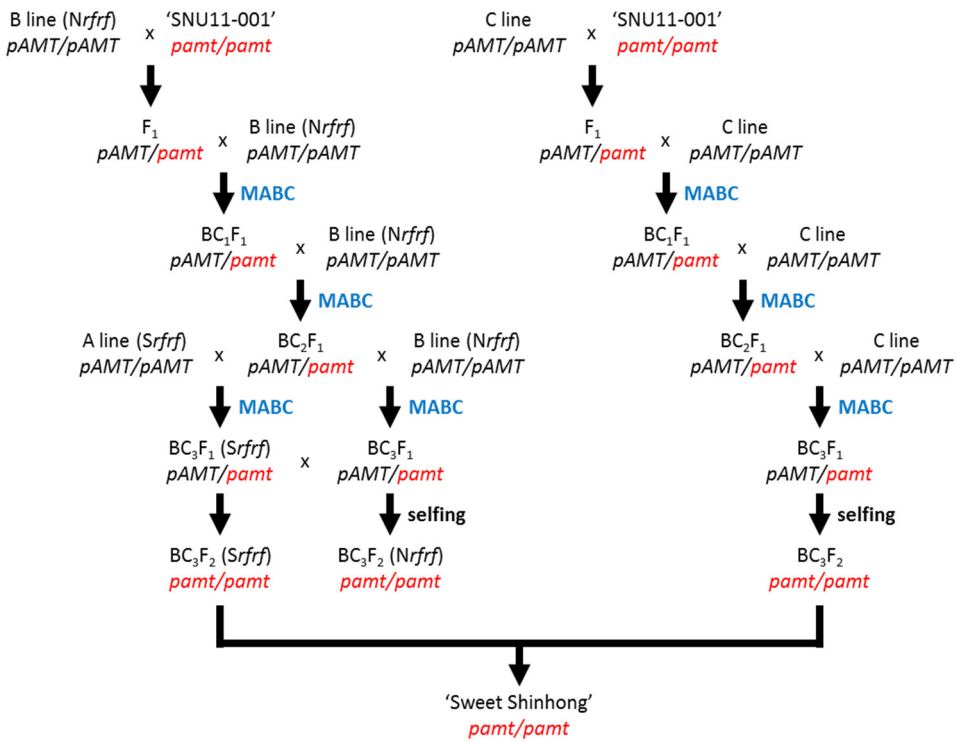

Fig. 3. Schematic diagram of breeding 'Sweet Shinhong'. *MABC* Marker-assisted backcrossing with foreground selection by using *pAMT* marker and background selection with EP1^{TM} system. *A line* CMS-line of 'Shinhong'. *B line* Maintainer of 'Shinhong'. *C line* Paternal-line of 'Shinhong'. *N* Normal cytoplasm. *S* Sterile cytoplasm. *rf* Recessive fertility restoring nuclear gene.

Table 5. Genotyping results of the foreground selection in ‘Sweet Shinhong C’ backcross population

Population	Generation	Number of plant			Expected ratio	χ^2	*P-value*
		Total	*pAMT/pAMT*	*pAMT/pamt*			
‘Shinhong C’ x ’SNU11-001’	BC_1F_1	250	109	141	1:1	4.0960	0.0430
	BC_2F_1	160	71	89	1:1	2.0250	0.1547

1.2 Background selection using the EP1^{TM} system

Plants selected by the *pAMT* mareker were subjected to background selection using EP1^{TM} system (Fig. 4A). A total of 141 BC_1F_1 plants were analyzed by 192 SNP markers and the recovery rates of the recurrent parent were calculated. The recovery rate was varied from 59.4 (BC_1F_1-63) to 84.7% (BC_1F_1-8) with an average recovery rate of 72.6% depending on plants. BC_1F_1-8 plant had a recovery rate of a 84.7 %, however, chromosome 1, 4, 9 and 11 still retained the 'SNU11-001' background (Fig. 5A). This plant was used for another round of backcrossing.

A total of 89 BC_1F_1 plants were screened 64 SNP markers as described above (Fig. 4B). The recovery rate was varied from 87.1 (BC_1F_1-63) to 99.1 % (BC_1F_1-8) with an average recovery rate of 92.5% depending on plants. Several plants were selected including 'BC_2F_1-40' (99.1%), 'BC_2F_1-8' (98.0%), 'BC_2F_1-36' (97.1%), 'BC_2F_1-43 (97.1%)', 'BC_2F_1-18' (96.9%), 'BC_2F_1-48' (96.9%) and 'BC_2F_1-46' (96.3%). 'BC_2F_1-40' plant recovered the 'Shinhong C' background in the most genome regions excluding 'SNU11-001' segment on chromosome 3 that contains the *pAMT* allele (Fig. 5B).

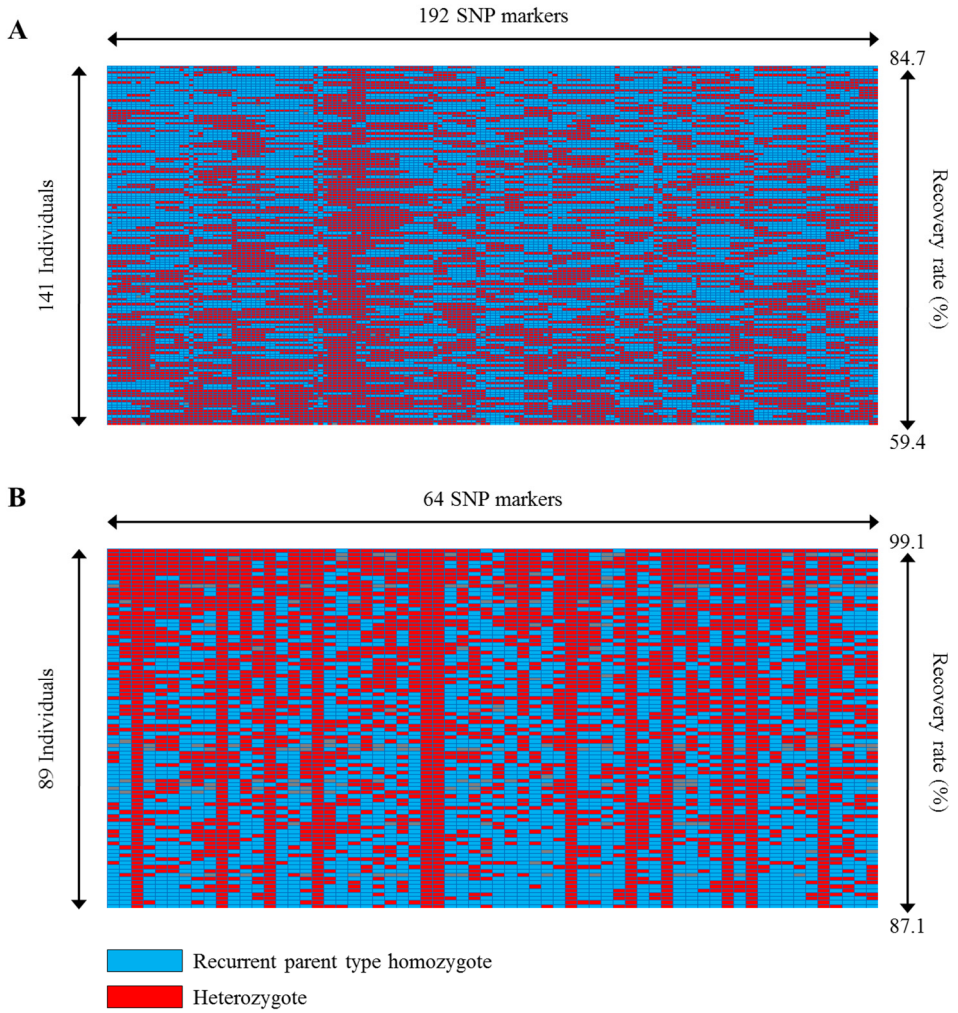

Fig. 4. Genotyping results of the background selection in backcross populations. **A** Genotyping results for BC_1F_1 population. A total of 141 individuals and 192 SNP markers were used. The highest recovery rate is 84.7% and the lowest recovery rate is 59.4%. **B** Genotyping results of BC_2F_1 population. A total of 89 individuals and 64 SNP markers were used. SNP markers homozygous in BC_1F_1 were not duplicated with markers used in BC_2F_1. The highest recovery rate was 99.1% and the lowest recovery rate is 87.1%. *Horizontal axis* shows genotypes of an individual plant. *Left vertical axis* shows individual plants used in this study, *Right vertical axis* shows the recovery rate (%) of each individuals. *Blue* recurrent parent type homozygote, *Red* heterozygote, *Gray* unknown.

Fig. 5. Graphical genotype of the selected plants, 'BC_1F_1-8' and 'BC_2F_1-40' in 'Shinhong C' backcross population. The red SNP markers on the chromosomes indicates homozygous region for the recipient genome while black SNP markers indicates heterozygote. The distances were represented in cM based on published map of Liu et al. (2014). **A** Graphical genotype of 'BC_1F_1-8' plant. 192 SNP markers were used for the background selection and 160 SNP markers were showed without 30 unknown genotypes. **B** Graphical genotype of 'BC_2F_1-40' plant. 64 SNP markers used in the BC_2F_1 generation were added to the **A** graphics.

Validation of MABC system in 'Shinhong C' backcross population using GBS

In order to validate MABC results, 'Shinhong C', 'SNU11-001' and selected BC_2F_1 plants were genotyped by the genotype-by-sequencing (GBS) method. A total 6,252 SNPs between 'Shinhong C' and 'SNU11-001' were detected by GBS (Table 6). SNPs were most abundant in chromosome 3 being 734 SNPs whereas only 338 SNPs were found in chromosome 8. The SNP density revealed by GBS was one SNP per every 448,509 nt.

The number of the SNP alleles of the recurrent parent was calculated for each selected BC_1F_1 and BC_2F_1 individual as the same way as EP1™ genotyping and results were compared (Table 7). BC_1F_1-8 plant had 4,135 homozygous alleles of 'Shinhong C', 1,604 heterozygous alleles and 155 homozygous alleles of 'SNU11-001'. When the recovery rate was calculated, BC_1F_1-8 showed 83.8%, which is similar to that was calculated by EP1™ genotyping. BC_2F_1-40 plant had 5,654 homozygous alleles of 'Shinhong C', 386 heterozygous alleles and 29 homozygous alleles of 'SNU11-001'. The recovery rate of BC_2F_1-40 was calculated to be 99.1% using GBS that is 2.8% lower than EP1™ genotyping result. Other selected BC_2F_1 plants also showed the less than 2.3% differences between GBS and EP1™ genotyping results. Overall, the recovery rates calculated by both methods were very similar and indicating a couple hundred SNP markers are enough for background selection.

Graphical genotype of the selected BC_2F_1 plants are shown in Fig. 6. The introgressed segments size of 'SNU11-001' allele on chromosome 3 are as follows: 'BC_2F_1-40'; 45,036,922 nt, 'BC_2F_1-8'; 153,157,307 nt, 'BC_2F_1-36'; 150,565,488 nt, 'BC_2F_1-43'; 111,997,267 nt, 'BC_2F_1-18'; 154,210,477 nt, 'BC_2F_1-48'; 106,119,677 nt and 'BC_2F_1-46'; 99,336,374 nt.

Table 6. SNPs polymorphic between 'Shinhong C' and 'SNU11-001'

Chromosome	Number of SNP	Chromosome size (nt)[a]	SNP density (nt)[b]
1	642	272,704,604	424,988
2	560	171,128,871	305,873
3	734	257,900,543	351,266
4	463	222,584,275	481,140
5	448	233,468,049	521,877
6	628	236,913,278	377,634
7	496	231,911,496	465,249
8	338	145,103,255	430,165
9	497	252,779,264	508,032
10	494	233,593,809	473,051
11	457	259,726,002	566,248
12	495	235,688,241	476,585
Total	6,252	2,753,501,687	448,509

[a] Chromosome size was obtained from CM334 20140122 (v1.55) CHROMOSOME database (http://peppergenome.snu.ac.kr).
[b] SNP density is average distance between SNPs.

Table 7. Comparison of the recovery rate between EP1™ genotyping and GBS

Sample name	Number of SNP			Recovery rate (%)[a]	
	Shinhong C	Heterozygote	SNU11-001	EP1™ genotyping	GBS
BC_1F_1-8	4,135	1,604	155	84.5	83.8
BC_2F_1-40	5,654	386	29	99.1	96.3
BC_2F_1-8	5,430	390	58	98.0	95.7
BC_2F_1-36	5,216	446	52	97.1	95.2
BC_2F_1-43	5,419	392	58	97.1	95.7
BC_2F_1-18	5,427	388	72	96.9	95.5
BC_2F_1-48	5,321	575	53	96.6	94.3
BC_2F_1-46	5,340	439	60	96.3	95.2

[a] Recovery rate of the 'Shinhong C' was calculated as the percentage of (number of 'Shinhong C' genotype + half number of the hetero genotype)/total number of genotype.

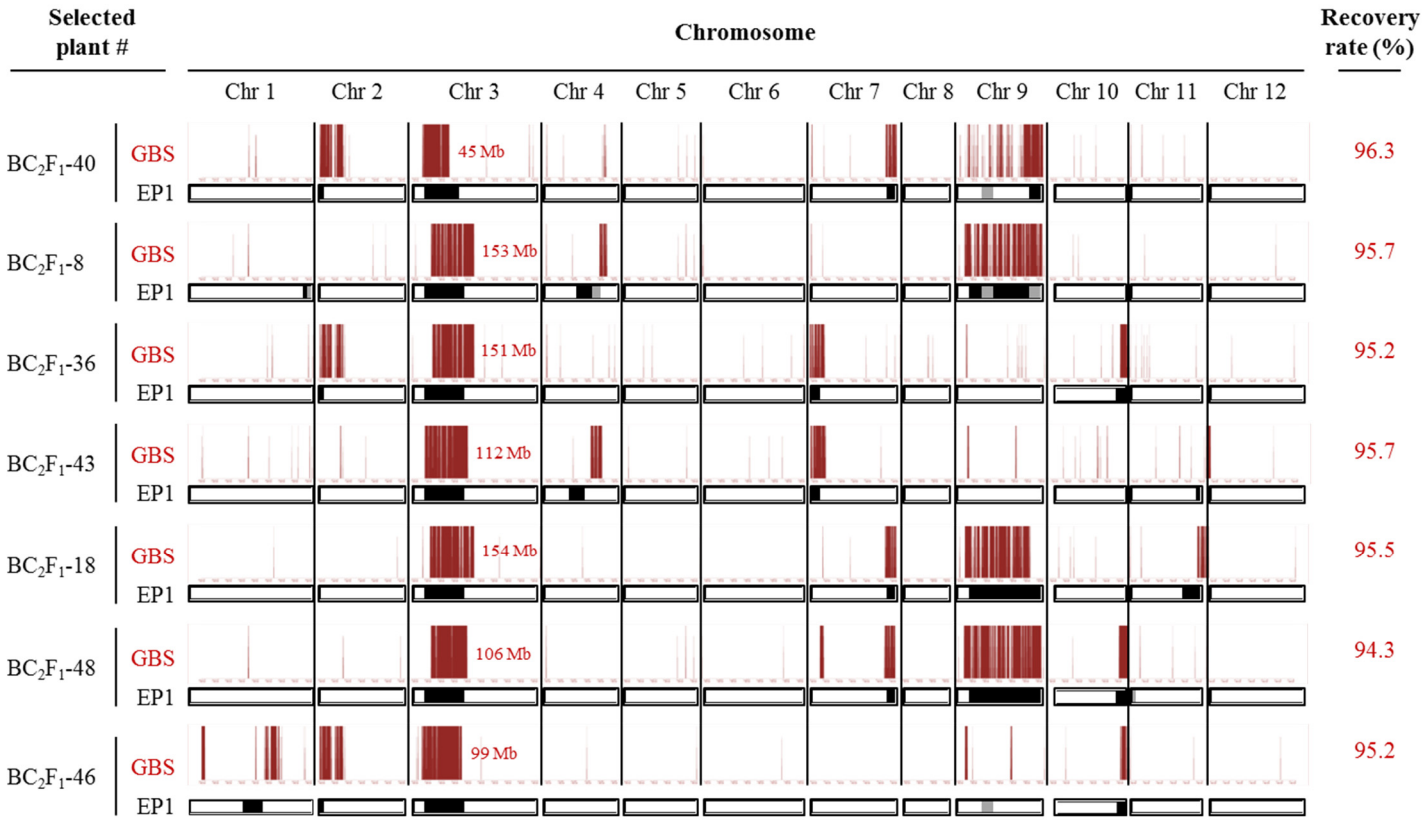

Fig. 6. Graphical genotype of the selected BC_2F_1 plants using GBS and comparison with EP1 results. *White region* means 'Shinhong C' genotype. *Red region* means heterozygote using GBS. *Black region* means heterozygote using EP1. *Gray region* means unknown genotype.

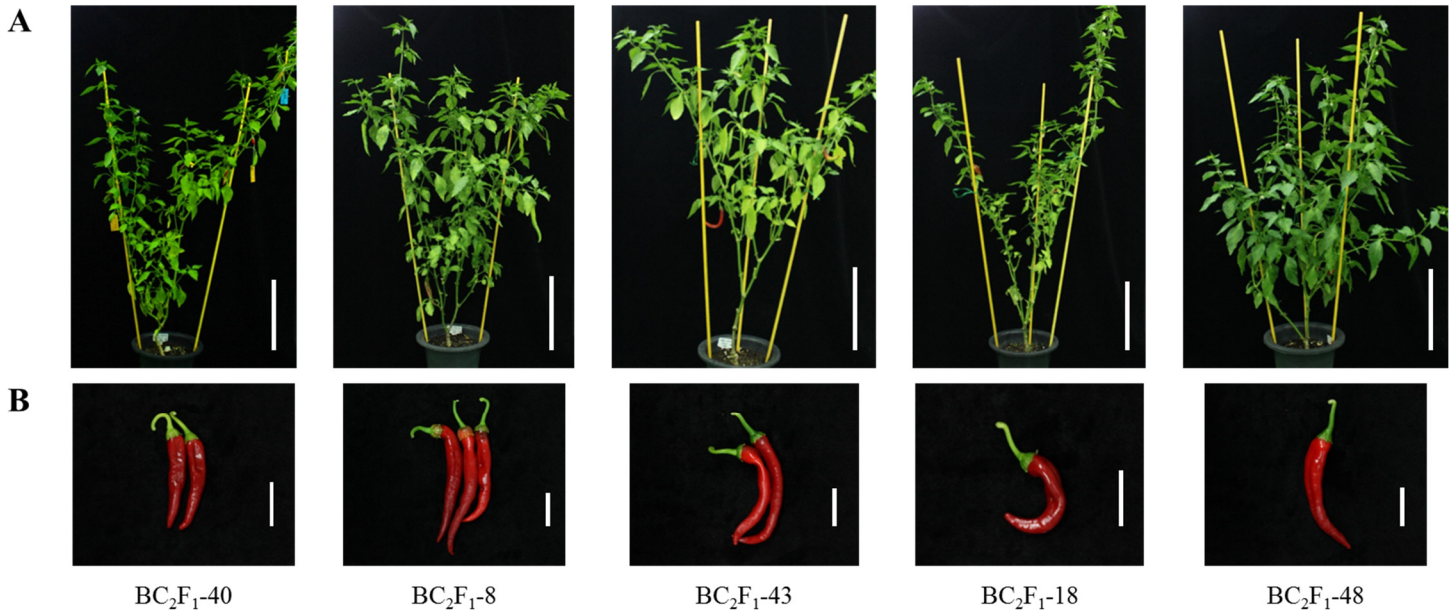

Fig. 7. Selected BC_2F_1 plants and their fruits. 'BC_2F_1-40', 'BC_2F_1-8', 'BC_2F_1-43', BC_2F_1-18' and 'BC_2F_1-48' is ordered from left to right. **A** Plants of the selected BC_2F_1 plants. *Bar* 11 cm. **B** Fruits of the selected BC_2F_1 plants. *Bar* 3 cm.

DISCUSSION

Biosynthesis of capsinoid is known to be caused by the functional loss of *pAMT* in *Capsicum* (Lang et al., 2009). Capsinoid has a similar physiological effects as capsaicinoid and has an advantage in that it has less pungent property (Sasahara et al., 2010). Previous studies showed that 'SNU11-001' synthesized a much higher amounts of capsinoid than other *Capsicum* spp. and the mutation of the *pAMT* gene is recessively inherited. The purpose of this study was introduction of the dysfuntional *pAMT* allle in *C. chinense* to a *C. annuum* commercial cultivar, 'Shinhong'. We used a marker-assisted backcross method to develop varieties containing high levels of capsinoid. To facilitate backcross breeding, a *pAMT* marker was developed and individuals containing the *pAMT* allele were selected using the marker. For background selection, SNP markers covering whole pepper genome were genotyped by a high throughput method.

As in other MABC program, our MABC program was also consisted of two steps, 'foreground selection' and 'background selection'. In the 'foreground selection' step, we used a SCAR marker which can discriminate the dysfunctional pAMT allele from the normal allele. In this step, heterozygotes of the *pAMT* locus and homozygote for the dominant *pAMT* allele were discriminated by the SCAR marker. The *pAMT* marker was developed using the *pAMT* sequence in 'Aji Dulce' strain 2 (Tanaka et al., 2010b) and that marker genotype was perfectly correlated

with the presence of capsinoid in 'Habanero' x 'SNU11-001' F_2 population (Jang et al., 2015).

The key in MABC is to select individuals plants most similar to the recurrent parent in each backcross generation. The selection of individual plants based on morphological traits is ineffective because it is time consuming and labor intensive. To overcome these limitations, molecular markers can be used for selection individuals having the most similar marker genotypes to the recurrent parent (Collard and Mackill, 2008). In this study, a total of 412 SNP markers distributed in whole pepper genome (Kang et al., 2014) were used to select a polymorphic marker set. Using 204 SNP polymorphic markers between the recurrent and donor parents, the background genotypes of backcross populations were determined. As a result, plants recovered 84.7% and 99.1% of the recurrent parent were selected in BC_1 BC_2 generations, respectively. Compared to the theoretical recovery rates, 75% in BC_1 and 87.5% in BC_2 generation, we were able to select individual plants with about 10% more recovery rate in out MABC program. This results were confirmed by genome-wide sequencing methods GBS. Therefore, the results showed that EP1^{TM} system is useful in MABC. In summary, GBS results showed that recovery rate and positions of introgressed segments were well matched between two methods. The average difference of recovery rate between the EP1^{TM} system and GBS was only about 1.7%. These results demonstrate that MABC can be successfully done with a couple hundred SNP markers.

REFERENCES

Collard B.C., Mackill D.J. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 363:557-572.

Curry J., Aluru M., Mendoza M., Nevarez J., Melendrez M., O'Connell M.A. 1999. Transcripts for possible capsaicinoid biosynthetic genes are differentially accumulated in pungent and non-pungent *Capsicum spp*. Plant Sci. 148:47-57.

Doyle J.J., Doyle J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. bull. 19:11-15.

Faraut B., Giannesini B., Matarazzo V., Le Fur Y., Rougon G., Cozzone P.J., Bendahan D. 2009. Capsiate administration results in an uncoupling protein-3 downregulation, an enhanced muscle oxidative capacity and a decreased abdominal fat content *in vivo*. Int. J. Obes. 33:1348-1355.

Frisch M., Bohn M., Melchinger A.E. 1999. Comparison of selection strategies for marker-assisted backcrossing of a gene. Crop Sci. 39:1295-1301.

Han K., Jeong H.J., Sung J.H., Keum Y.S., Cho M.C., Kim J.H., Kwon J.K., Kim B.D., Kang B.C. 2012. Biosynthesis of capsinoid is controlled by the *Pun1* locus in pepper. Mol. Breed. 31:537-548.

Herzog E., Frisch M. 2011. Selection strategies for marker-assisted backcrossing with high-throughput marker systems. Theor. Appl. Genet. 123:251-260.

Hospital F., Charcosset A. 1997. Marker-Assisted Introgression of Quantitative Trait Loci. Genetics 147:1469-1485.

Hospital F., Chevalet C., Mulsant P. 1992. Using Markers in Gene Introgression Breeding Programs. Genetics 132:1199-1210.

Iida T., Moriyama T., Kobata K., Morita A., Murayama N., Hashizume S., Fushiki T., Yazawa S., Watanabe T., Tominaga M. 2003. TRPV1 activation and induction of nociceptive response by a non-pungent capsaicin-like compound, capsiate. Neuropharmacology 44:958-967.

Jang S., Han K., Jo Y.D., Jeong H.J., Siddique M.I., Kang B.C. 2015. Substitution of a Dysfunctional *pAMT* Allele Results in Low-Pungency but High Levels of Capsinoid in *Capsicum chinense* 'Habanero'. Plant Breed. Biotechnol. 3:119-128.

Johnston T.D. 1974. Transfer of disease resistance from *Brassica campestris* L. to rape (*B. napus* L.). Euphytica 23:681-683.

Kang J.H., Yang H.B., Jeong H.S., Choe P., Kwon J.K., Kang B.C. 2014.

Single Nucleotide Polymorphism Marker Discovery from Transcriptome Sequencing for Marker-assisted Backcrossing in *Capsicum*. Korean J. Hortic. Sci. 32:535-543.

Kawabata F., Inoue N., Masamoto Y., Matsumura S., Kimura W., Kadowaki M., Higashi T., Tominaga M., Inoue K., Fushiki T. 2009. Non-pungent capsaicin analogs (capsinoids) increase metabolic rate and enhance thermogenesis *via* gastrointestinal TRPV1 in mice. Biosci. Biotechnol. Biochem. 73:2690-2697.

Kobata K., Sugawara M., Mimura M., Yazawa S., Watanabe T. 2013. Potent production of capsaicinoids and capsinoids by *Capsicum* peppers. J. Agric. Food Chem. 61:11127-11132.

Kobata K., Sutoh K., Todo T., Yazawa S., Iwai K., Watanabe T. 1999. Nordihydrocapsiate, a new capsinoid from the fruits of a nonpungent pepper, *Capsicum annuum*. J. Nat. Prod. 62:335-336.

Kobata K., Todo T., Yazawa S., Iwai K., Watanabe T. 1998. Novel capsaicinoid-like substances, capsiate and dihydrocapsiate, from the fruits of a nonpungent cultivar, CH-19 sweet, of pepper (*Capsicum annuum* L.). J. Agric. Food Chem. 46:1695-1697.

Lang Y., Kisaka H., Sugiyama R., Nomura K., Morita A., Watanabe T., Tanaka Y., Yazawa S., Miwa T. 2009. Functional loss of *pAMT* results in biosynthesis of capsinoids, capsaicinoid analogs, in *Capsicum annuum* cv. CH-19 Sweet. Plant J. 59:953-961.

Liu W.Y., Kang J.H., Jeong H.S., Choi H.J., Yang H.B., Kim K.T., Choi D., Choi G.J., Jahn M., Kang B.C. 2014. Combined use of bulked segregant analysis and microarrays reveals SNP markers pinpointing a major QTL for resistance to *Phytophthora capsici* in pepper. Theor. Appl. Genet. 127:2503-2513.

Luo X.J., Peng J., Li Y.J. 2011. Recent advances in the study on capsaicinoids and capsinoids. Eur. J. Pharmacol. 650:1-7.

Park Y.J., Nishikawa T., Minami M., Nemoto K., Iwasaki T., Matsushima K. 2015. A low-pungency S3212 genotype of *Capsicum frutescens* caused by a mutation in the *putative aminotransferase* (*p-AMT*) gene. Mol. Genet. Genomics:1-8.

Pyun B.J., Choi S., Lee Y., Kim T.W., Min J.K., Kim Y., Kim B.D., Kim J.H., Kim T.Y., Kim Y.M., Kwon Y.G. 2008. Capsiate, a nonpungent capsaicin-like compound, inhibits angiogenesis and vascular permeability via a direct inhibition of Src kinase activity. Cancer Res. 68:227-235.

Ragot M., Biasiolli M., Delbut M., Dell'Orco A., Malgarini L., Thevenin P., Vernoy J., Vivant J., Zimmermann R., Gay G. 1995. Marker-assisted

backcrossing: a practical example. COLLOQUES-INRA:45-45.
Randhawa H.S., Mutti J.S., Kidwell K., Morris C.F., Chen X., Gill K.S. 2009. Rapid and targeted introgression of genes into popular wheat cultivars using marker-assisted background selection. PLoS One 4:e5752.
Sasahara I., Furuhata Y., Iwasaki Y., Inoue N., Sato H., Watanabe T., Takahashi M. 2010. Assessment of the Biological Similarity of Three Capsaicin Analogs (Capsinoids) Found in Non-Pungent Chili Pepper (CH-19 Sweet) Fruits. Biosci. Biotechnol. Biochem. 74:274-278.
Stoskopf N.C., Tomes D.T., Christie B.R. (1993) Plant breeding: Theory and practice. Westview Press Boulder
Sutoh K., Kobata K., Watanabe T. 2001. Stability of capsinoid in various solvents. J. Agric. Food Chem. 49:4026-4030.
Tanaka Y., Hosokawa M., Miwa T., Watanabe T., Yazawa S. 2010a. Newly mutated *putative-aminotransferase* in nonpungent pepper (*Capsicum annuum*) results in biosynthesis of capsinoids, capsaicinoid analogues. J. Agric. Food Chem. 58:1761-1767.
Tanaka Y., Hosokawa M., Miwa T., Watanabe T., Yazawa S. 2010b. Novel loss-of-function *putative aminotransferase* alleles cause biosynthesis of capsinoids, nonpungent capsaicinoid analogues, in mildly pungent chili peppers (*Capsicum chinense*). J. Agric. Food Chem. 58:11762-11767.
Tanaka Y., Yoneda H., Hosokawa M., Miwa T., Yazawa S. 2014. Application of marker-assisted selection in breeding of a new fresh pepper cultivar (*Capsicum annuum*) containing capsinoids, low-pungent capsaicinoid analogs. Sci. Hortic. 165:242-245.
Tanksley S., Rick C. 1980. Isozymic gene linkage map of the tomato: applications in genetics and breeding. Theor. Appl. Genet. 58:161-170.
Tanksley S.D. 1983. Molecular markers in plant breeding. Plant Mol. Biol. Rep. 1:3-8.
Thiele R., Mueller-Seitz E., Petz M. 2008. Chili pepper fruits: presumed precursors of fatty acids characteristic for capsaicinoids. J. Agric. Food Chem. 56:4219-4224.
Voorrips R. 2002. MapChart: software for the graphical presentation of linkage maps and QTLs. J. Hered. 93:77-78.
Yazawa S., Suetome N., Okamoto K., Namiki T. 1989. Content of capsaicinoids and capsaicinoid-like substances in fruit of pepper (*Capsicum annuum* L.) hybrids made with" CH-19 Sweet" as a Parent. J. Jpn. Soc. Hortic. Sci. 58:601-607.

초 록

캡시노이드는 캡사이시노이드와 달리 매운 맛이 약한 물질이지만 캡사이시노이드처럼 항암 효과나 항비만 효과와 같은 생리적 작용을 하는 물질이다. 캡시노이드의 전구체인 바닐릴 알코올은 *putative-aminotransferase* (*pAMT*)의 돌연변이에 의해 생성된다고 알려져 있다. 선행연구를 통해 *Capsicum* 유전자원에서 캡시노이드 고 함유 계통인 'SNU11-001' (*Capsicum chinense*)를 발견하였다. 'SNU11-001'에서 *pAMT* 유전자가 기능을 하지 못하도록 하는 돌연변이가 있음을 확인하였다. 캡시노이드 고함유 고추 품종을 개발하기 위해 여교배 육종 과정에서 분자표지를 활용한 foreground selection과 background selection을 수행하였다. 전통적인 여교배 육종법과 비교하였을 때, 분자표지를 활용한 여교배 육종법(MABC)은 반복친의 유전형을 빨리 회복시키는데 매우 효과적이다. Foreground selection 과정에서 'SNU11-001'의 *pAMT* 돌연변이를 탐지할 수 있는 SCAR 분자표지를 사용하여 BC_1F_1과 BC_2F_1 집단에서 *pAMT/pamt* 유전형을 가진 개체들을 선발하였다. Background selection 분자표지를 개발하기 위하여 412개의 단일염기다형성(SNP) 분자표지를 사용하여 '신홍'의

부모친 계통과 ‘SNU11-001’ 사이에 다형성 검정을 실시하였다. 412개의 단일염기다형성 분자표지 중 고추 유전체에 골고루 분포한 144개의 분자표지와 204개의 분자표지를 최종적으로 선발하였다. 선발된 분자표지를 사용하여 *pAMT/pamt* 유전형을 가지고 있는 BC_1F_1과 BC2F1 개체들을 대상으로 background selection을 수행하였다. 대량 유전형 분석 시스템 (EP1™, Fluidigm®, USA)를 사용하여 다량의 유전형을 분석하였다. Background selection 결과 반복친 유전형을 84% 회복한 BC_1F_1개체와 96% 이상 회복한 BC_2F_1 개체들을 선발하였다. 단일염기다형성 분자표지를 활용한 background selection의 결과를 확인하기 위하여 genotype-by-sequencing (GBS) 방법을 사용하여 선발된 BC_2F_1 개체들의 유전형을 분석하였다. GBS의 결과를 통해 반복친 유전형 회복률과 ‘SNU11-001’의 유전형 위치가 MABC의 결과와 상당히 일치함을 확인하였고 따라서 수백개의 단일염기다형성 분자표지를 활용한 MABC가 성공적으로 수행되었음을 보여준다.

주요어: 캡시노이드, *pAMT*, marker-assisted backcross breeding

학번: 2013-23232